

#2

520.39555X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE



Applicant(s): SHIMA, ET AL.
Serial No.:
Filed: January 26, 2001
Title: DATA MANAGEMENT SYSTEM FOR STORAGES
Group:

LETTER CLAIMING RIGHT OF PRIORITY

Honorable Commissioner of
Patents and Trademarks
Washington, D.C. 20231

January 26, 2001

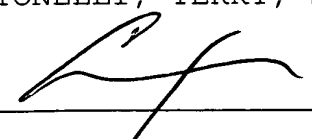
Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Patent Application No.(s) 2000-197861 filed June 27, 2000.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Carl I. Brundidge
Registration No. 29,621

CIB/mdt
Attachment
(703)312-6600

BEST AVAILABLE COPY

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT

JCS29 U.S. PTO
09/769270
01/26/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application:

2000年 6月27日

出 願 番 号
Application Number:

特願2000-197861

出 願 人
Applicant(s):

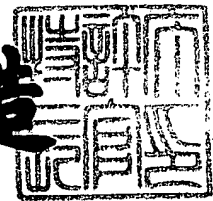
株式会社日立製作所

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年10月 6日

特許庁長官
Commissioner,
Patent Office

及 川 耕 造



出証番号 出証特2000-3081226

【書類名】 特許願

【整理番号】 NT00P0191

【提出日】 平成12年 6月27日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 12/00

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所 ストレージシステム事業部内

【氏名】 島 誠二

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所 ストレージシステム事業部内

【氏名】 村上 正治

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所 ストレージシステム事業部内

【氏名】 光岡 芳夫

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100068504

【弁理士】

【氏名又は名称】 小川 勝男

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100086656

【弁理士】

【氏名又は名称】 田中 恭助

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100094352

【弁理士】

【氏名又は名称】 佐々木 孝

【電話番号】 03-3661-0071

【手数料の表示】

【予納台帳番号】 081423

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 記憶装置のデータ管理システム

【特許請求の範囲】

【請求項 1】

複数の記憶装置がデータ転送網に接続されたシステムにおいて、異なった OS に固有の意味のあるデータの塊を、前記データ転送網に共通の意味を持つデータの単位に変換する変換機構と、ホストから分離されて設けられ前記ホストから前記データの単位名を受けて前記記憶装置の 1 つからの前記データの単位の読み出しを管理する機構とを備えたことを特徴とする記憶装置のデータ管理システム。

【請求項 2】

複数の記憶装置とホストがデータ転送網に接続されたシステムにおいて、前記記憶装置からファイルを取得するホストと、前記ホストと分離して存在するファイルを管理するサーバと、前記ホストの持つ異なった OS に固有のフォーマットのファイルを前記データ転送網に共通の意味を持つフォーマットを持つ共通フォーマットファイルに変換する変換機構とを有し、前記サーバは前記共通フォーマットファイル名で前記ホストからファイルのアクセス許可要求を受けると前記記憶装置の前記ファイルのホストへの送信を管理することを特徴とする記憶装置のデータ管理システム。

【請求項 3】 前記サーバは記憶装置に対してファイル操作 ID を前記アクセス許可要求を受けたファイルに付してステージングを要求し、かつホストに前記ファイル操作 ID を送信し、前記ホストは前記ファイル操作 ID を付して前記記憶装置にファイル操作要求を出して当該ファイルを取得することを特徴とする請求項 2 記載の記憶装置のデータ管理システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、異なったフォーマットのデータが混在する計算機システムにおけるファイル記憶管理システムに関する。

【 0 0 0 2 】

【従来の技術】

磁気ディスク装置の様な書換え可能な記憶媒体を用いた記憶装置は従来、ホストと1対1の関係に有り、OSにその間のデータのやり取りの管理を依存するという構成となっている。ところで、記憶媒体は年々、記憶容量の増加、ダウンサイジング化が進み、その結果、大容量ボリュームを複数個持つ構成のディスクアレー装置が出現し、更に技術が進んだ現在は、信頼性/冗長性を持ったRAID装置が実現されている。この記憶装置は他の装置との接続口であるポートを複数個持ち、其々のポートにホストが接続可能な、ホストとは1対多の接続関係になってきている。ところが、ホストからの記憶媒体を認識する技術は変わらず、ホストから記憶装置へのパスを指定し1対1でデータのやり取りをする方式を取っている。その結果、RAID装置内に互換性の無いフォーマットのデータが混在し、管理をより複雑なものにしている。

【 0 0 0 3 】

更に、このような状況において、近年SAN (Storage Area Network) といったストレージのデータ転送の高速化を目的とした技術が普及してきている。これは、ファイバ (Fibre) ネット網により多数のホストと記憶装置が接続され、LANに比べて遥かに高速のデータの通信が出来ると共に、ホストの関与なく記憶装置間でデータの高速転送が可能なものである。

【 0 0 0 4 】

公知例としては例えば、ファイル変換に関するものとして特開平11-134227号公報がある。これは、あるOSのあるファイルシステムから他のOSのあるファイルシステムにファイルフォーマットを変換するものである。これから判るようにこの公知例は記憶装置がSANに接続されているような新しい環境に適合することは考慮されていない。

【 0 0 0 5 】

【発明が解決しようとする課題】

以上に述べた様に、従来の技術は複数のポートを持った記憶装置を異なったOS (独自のファイルシステムを有する) が共有している。各OSのファイルシス

テムは独自のファイルフォーマットでペイロードされた記憶装置内では意味の異なったデータの塊で直接記憶媒体にリード／ライトして目的のファイルデータを取得している為、互換性の無いデータが混在している。

【0006】

この状態では、記憶装置は、“意味の異なったOS単位のデータの塊”か“記憶媒体の集合”を単位としてデータを管理することしか出来ず、記憶装置内でボリュームを自動的に拡張したりする等のストレージ管理はいつでもOSの管理の下で実現するしかない。

【0007】

本発明の目的は、SANの様な高速データ転送技術を持ったシステムに於いて、記憶装置内で意味を持った塊（ファイル単位等）でデータを管理し、OSの種類を意識することなく、理論的には無制限な容量をユーザーに提供できる広範囲なデータ管理を実現することにある。

【0008】

【課題を解決するための手段】

本発明は、ホストに固有なフォーマットを持ったデータの意味のある塊（ファイル単位）を、複数の記憶装置間で共通なフォーマットに変換する機能を有し、更に、これらの複数の記憶装置を制御するサーバをホストとは別に備えたものであり、これらの記憶装置に共通のファイルシステムを構成し、記憶装置間で共通の塊（ファイル）でアクセス可能とすることを特徴とする。

【0009】

【発明の実施の形態】

以下、発明の実施形態について図面を用いて説明する。

【0010】

図1で、ホスト-A（A-1）、ホスト-B（A-2）、SAN-FM（ファイルマネージャ）（A-3）、SAN-M（マネージャ）（A-4）、RAID-A（A-5）、RAID-B（A-6）はLAN（A-8）でそれぞれ接続線L1からL6によって接続されている。又、これらは、ファイバスイッチ（A-7）経由でSAN（A-9）に接続ポートS1からS8によって接続されている

。ホストA、Bはワークステーション、SAN-FS、SAN-Mはそれぞれサーバーマネージャステーション、RAID-A、BはFibre I/Fを持ったRAID装置である。ホスト-A、B、SAN-FS、SAN-Mは違ったOSで動作している。ホスト-A、Bは、RAID-A、B内の論理デバイス（論理ボリューム）にアクセスすることが可能で、RAID-A、Bは、独立したIDを持ったFibreポートを複数（5～8）持っていて、その個々のポートはSAN（A-9）に接続できる。

【0011】

図1は1つのファイバスイッチに接続された装置を示している。このファイバスイッチは、図示していないがSAN（A-9）を通して他のファイバスイッチに接続されている。従って、膨大な数のRAIDが互いにSANを通じて接続され得る。

【0012】

図1に示すように、ホストとは分離したSAN-FM、SAN-Mが設けられている。これにより記憶装置をホストから独立させホストのファイル形式に関係なく記憶装置間での自由なデータの転送を可能としたものである。これが本実施例において特徴の一つとするところであり、その構成、動作については後述する。なお、ここでRAIDとは記憶部のドライブ単位で冗長性を持ち、複数のポートを持つ記憶装置を指す。本発明ではRAIDに限らずホストの外にある外部記憶装置であれば良い。ホストは記憶装置にとって上位装置に当たる。

【0013】

図2で、RAID内には、複数の物理デバイス（B-1）（例えばディスクドライブ）で構成されている物理デバイス群（B-2）（物理ボリューム群又はECCグループと呼ぶ）が複数存在する。この物理ボリュームは良く知られているように論理的に分割され、任意のサイズの複数の論理ボリューム（B-4）が構築される。この論理ボリュームは個々にLUN（Logical Unit Number：ディスク-ホスト間のパスを定義するもの）を持つ。論理ボリュームはポートに接続が可能で、ホストのWWnとポートのWWnと論理ボリュームが指定されればホストとRAIDのデータ転送のパスの条件は揃う。論理ボリュームの構

成情報は、RAID内のRAID-DBにシステム管理情報の一部として管理される。

【0014】

例えば、図2の様な構成を持ったRAIDがあった場合、ホストは、ポート0 (S5) からLUN0にアクセスをすると論理ボリュームA(B-4)を操作(データの更新: 更新、訂正、削除、追加)することが可能である。個々のポートに任意のWWnの名前が付加されているとすると、ホストは、ルート情報(自分自身のWWn、RAIDのWWn、ポート0 (S5)、LUN0)を指定して論理ボリュームAを操作することができる。

【0015】

図3は本発明のファイルの変換を模式的に表したものである。なお、ファイルとは前述した意味のあるデータの塊の単位の代表的なものであり、以下に、ファイルを例にとって説明するが別の単位であっても良い。一般にファイル(1, 2, 3)は管理部(4, 5, 6)と情報部(7, 8, 9)に分かれている。管理部(4, 5, 6)はそのファイル定義の形式がホスト毎に異なると共に、図のように管理部の付与形式も多様である。即ち、ホストに固有なファイル形式となっている。本発明では、あたかもホストに固有なファイルに被せるが如くに共通な形式のファイル管理部(10, 11, 12)を管理部(4, 5, 6)から必要な情報を取り出しながら作る。そして、元のホスト固有のファイル(1, 2, 3)は全体がその情報部(13, 14, 15)として取り扱われる。

【0016】

なお、本実施例では情報部(13, 14, 15)は後述するSAN参加ID(K-1)で暗号化される。即ち、変換(24, 25, 26)はホスト固有なファイルに対して共通な管理部を作成することと、情報部の暗号化を意味する。

【0017】

次に図4によってSANで共通に取り扱われる論理ボリュームとそのLUNの構造を示す。ファイル(30, 31, 32)はそれぞれ適当なLUNに格納される。LUNはSANファイルディレクトリ(C-1)とファイルエリア(C-2)に分かれている。あるLUNに格納されたファイル(例えば30)の管理部(

例えば10)から必要な情報が選択され、SANファイルディレクトリ(C-1)に格納され、ファイル(30)がファイルエリア(C-2)に格納される。ファイルエリア(C-2)に格納されたファイルはSANファイルディレクトリ(C-1)から取得できるようにリンクが張られる。

【0018】

論理ボリューム(B-4)はSANファイルディレクトリ(C-1)とファイルエリア(C-2)から構成されていて複数のファイルを格納している。SANファイルディレクトリ(C-1)はファイルディレクトリのような階層構造でシステム管理情報(E-5)等を管理している。(この階層構造はミラーリング等の信頼性の都合でファイルやディレクトリが重複することが有り得る。)

ファイルエリア(C-2)には、実際のファイルデータが格納されている。SANファイルディレクトリ(C-1)内には、複数のボリュームラベル(C-1-1)が存在し、多重な管理が可能となっている。そのボリュームラベル(C-1-1)は、Holder Name(C-1-10)と1対多の関係を持っている。更に、Holder Name(C-1-10)は固有の属性と名前を持つことが可能で、File Name(C-1-11)と1対多の関係を持っている。File Name(C-1-11)は固有の属性と名前を持つことが可能で、所有者名(C-1-2)、利用者識別名(C-1-3)、WWn所有者名(C-1-4)、SCSI所有者名(C-1-5)、作成日時(C-1-6)、最終変更日時間(C-1-7)、最終参照日時(C-1-8)、ボリュームアトリビュート(C-1-9)等の属性を持っている。

【0019】

ボリュームアトリビュート(C-1-9)には、SAN上のボリュームに関する情報を記録することが可能である。所有者名(C-1-2)には、SAN上の管理情報(セキュリティ等)を記録することが可能である。利用者識別名(C-1-3)には、SAN上の管理情報(セキュリティ等)を記録することが可能である。WWn所有者(C-1-4)には、LUN自身を管理している接続先(ホスト側)WWnと接続元(RAID側)WWnの情報を記録することが可能である。SCSI所有者(C-1-5)には、LUN自身を管理しているSCSIパス情報を

記録することが可能である。(接続先/接続元)作成日(C-1-6)には、ボリュームを生成した日時を記録することが可能である。最終変更日時(C-1-7)には、ボリューム内のファイルを最後に変更した日時を記録することが可能である。最終参照日時(C-1-8)には、ボリューム内のファイルを最後に参照した日時を記録することが可能である。

【0020】

更に、File Name (C-1-11)は個々のファイルに対してFile属性(C-1-13)と、Data格納領域情報(C-1-12)等の属性を持っている。File属性(C-1-13)には、グローバル・オーナー(C-1-15)と、グローバル・グループ・オーナー(C-1-16)と、ファイルアトリビュート(C-1-14)を持っている。グローバル・オーナー(C-1-15)には、操作するオペレータ(人間)が任意に決定したオーナー名(パスワード)を記録し、グローバル・グループ・オーナー(C-1-16)には、操作するオペレータ(人間)が任意に決定したオーナーグループ名(パスワード)を記録し、ファイルアトリビュート(C-1-14)は、オペレータがファイル操作する時に必要になる個々のファイルの詳細なFile属性を記録することが可能で、OS-AP-FS等の固有の情報を記録することが可能である。Data格納領域情報(C-1-12)には、ファイルエリア内に格納されたファイルの情報(ファイルポインタ、ファイルデータ格納アドレス)を持っていて、ファイルエリア内に格納されたデータがどのような状態で作成されたものかを知ることが出来る。

【0021】

図5はLUNからFile Nameをキーとして抜き出した情報である。これはファイル管理のためにホストのG-DB内に格納しておく。これは目的のファイル名を指定し、そのパス情報を求めて、ホスト-RAID間で論理ボリュームにアクセスするためのものである。

【0022】

図6はRAID内での管理に使用されるRAID-DBを示す図である。RAID-DBとはパスの情報を管理するもので、ボリュームの構成情報、ボリュー

ムの障害情報を記憶するものである。図で、RAID固有名(E-2)は、固有の属性を持つことが可能であり、更に、唯一の存在である固有な名前であり、そしてWWn所有者名(E-3)と1対多の関係を持っている。WWn所有者名(E-3)は、固有の属性を持つことが可能であり、そして、任意の名前を持っていて、(しかし、WWn所有者名を重複させて、ポート単位にミラーリングをすることができる。)更に、ポートと1対1の関係を持っていて、そして、LUN所有者名(E-4)とは1対多の関係を持っている。LUN所有者名には、システム管理情報(E-5)、システム障害情報(E-6)等の管理情報を格納することが可能である。システム管理情報(E-5)には、RAID内のLUNのSANファイルディレクトリ等の情報を記憶する(RAIDに依存した物理デバイス構成情報も記録してある)。システム障害情報(E-6)には、PIN、縮退、閉塞等の危機/障害等の情報を記憶する。

【0023】

図7は図1の詳細を示す図である。図8は図7の動作を示しており、ホスト固有のファイルをSAN形式ファイルに変換してRAIDに格納する状態を示したものである。以下図7、図8を合わせて説明する。図8においてホストA、ホストB、ホストCに図7のそれぞれのFSによって違ったファイルフォーマットで作成されたそれぞれ異なるホスト固有のフォーマットを持つファイルG-1、G-2、G-3が存在する。図8で丸、三角、四角は単に異なるファイルであることを示したものである。これを図7のSAN-FSによって同一LUNに格納/管理できる様に型を変換する(G-4、G-5、G-6に)。図8で多角形で囲んであるのはSAN形式のファイルフォーマットであることを表している。

【0024】

図3で説明したように、SAN-FSはファイルフォーマットをホスト固有の管理部をSANに共通なものに変えるのではなく、SAN用のファイルフォーマットを被せる方法でファイルフォーマットを変換した後LUNに格納する機能を持つ。更に、変換の際に固有の鍵でファイルを暗号化する機能を持っている。又、逆にLUNに格納されたファイル(G-4、G-5、G-6)を読み出し、元のファイルフォーマット(G-1、G-2、G-3)に変換する機能を持っている(

LUNへの読み書きは一般的なファイルシステム（FS）のと同じ様に、変更/参照によりファイルエリアに格納されたファイル进行操作することができる）。STR-C（Storage Controller：F-3、F-4）間でのデータ操作で、第三者（クラッカー）からの介入操作を、転送（発信/着信）時間、独自暗号化フォーマット、ファイル操作履歴等でネットチェックすることが可能となっている。

【0025】

STR-C（F-3、F-4）は、ホストから受信したファイル、LUNから読み込んだファイルを非同期に仮想空間（G-7）に割当てる（ステージング）機能を持つ。また、非同期にファイルをホストに送信することと、LUN（G-8）にファイルを格納する（デステージング）機能を持つ。STR-Cは仮想空間（G-7）から実空間（G-8）へ、実空間（G-8）から仮想空間（G-7）への変換制御/管理/スペース確保を自動的に行い、資源の有効活用をスムーズに行う機能を持つ。その際、LUN内のSANファイルデレクトリのデータも更新することが可能である。またSTR-CはSAN-FSの様なファイルシステムの機能も持っている。STR-Cによれば、ホストが一つのトランザクションに占有されることなく、RAID間でファイル操作をすることが可能である。STR-Cはこれらの情報をリアルタイムにRAID-DBに格納され管理する機能を持つ。RAID-DBの更新がされた場合は、SAN-FS、SAN-MにDB（FS-DB、M-DB）の更新を要求する機能を持っている。又、SAN-FM、SAN-MにRAID-DBの更新情報を種類別で臨機応変に連絡（送信）する機能を持っていて、危機/障害が発生した時に、リアルタイムにSAN-MとSAN-FSに危機/障害情報を種類別で臨機応変に連絡（発信）することが可能である。又、その詳細な危機/障害情報をSAN-FMとSAN-Mに種類別で臨機応変に連絡（送信）することができる。

【0026】

SAN-Mサーバは、STR-Cを通して、RAID-DBの管理情報を取得と書換え（リフレッシュ）を要求する機能を持つ。STR-Cからの要求で、SAN-Mサーバは、STR-Cを通してRAID-DBの管理情報の取得と書換

え（リフレッシュ）を要求する。SAN-Mサーバは、RAID-DBで管理されたシステム構成情報やシステム管理情報を取得し、M-DBで管理することが可能である。SAN-MサーバはSAN-FSサーバにFM-DBの更新を要求する機能を持ち、システム管理情報を送信する機能を持っている。SAN-MサーバはSTR-Cがリアルタイムに発信する危機/障害情報を受信することが可能で、ホストとSAN-FMに障害情報を場合によりリアルタイムに連絡する機能を持っている。また、SAN-MサーバはホストかSAN-FMの要求で障害情報を送信することが可能で、RAID内の論理ボリュームを生成する機能を持っている（論理ボリューム構築はシステム管理情報の一部）。SAN-Mサーバは、一つのSAN環境グループの一括した管理をする。

【0027】

SAN-FSサーバは、ホストがSAN-FMサーバにアクセスして管理エリアのファイル进行操作（作成、更新、削除、参照等）することが可能である。SAN-FSサーバはSAN-Mからシステム管理情報を取得し、FS-DBで管理することが可能である。SAN-FSサーバはシステム管理情報をキーとして、STR-Cを通して詳細なSANファイルデレクトリ情報を読み書きする機能を持っていて、それで取得した情報は、FS-DBで管理する機能を持っている。

SAN-FSサーバはSTR-Cがリアルタイムに発信する危機/障害情報を受信することが可能である。SAN-FSサーバはホストにその障害情報を場合によりリアルタイムに連絡する機能を持っている。SAN-FSサーバは一つのSAN環境グループ内のファイルデータをファイル名で一括管理している。

【0028】

Driverには、FiberドライバやSCSIドライバやioctl（I/O Control）ドライバやFiberドライバが有り、又、それらをサポートする下層ドライバや上層ドライバが存在し、ファイルデータをSAN経由でRAIDから取得する機能を持ちSAN-FSの命令で機能する。

【0029】

NET-M（F-9～F-13）は、標準的なトランスポート・プロトコルでシステム管理情報、システム障害情報を転送する機能を持っている。

【 0 0 3 0 】

SAN (Storage Area Network) は、LUN内の情報 (詳細なSANファイルディレクトリ(C-1)、ファイルエリア(C-2))を転送するネットワークとして利用される。図示していないがDBMS (Data Base Management System: データベース管理システム) がRAID-DB、FM-DB、M-DB、G-DBのデータベースを管理する。ファイルシステム (FS) は、各OSで違ったファイルフォーマットを持っている。

【 0 0 3 1 】

図8に戻って、G-7に転送された変換後の各ファイルは、RAID-Aの複数個あるLUNの何れかに格納され得るし、また、SANで接続され他のファイバスイッチに接続された例えばRAID-CのLUNにも格納される。G-7に格納されたファイルはSANに接続されたRAID間で自由にやり取りが出来る。

【 0 0 3 2 】

次に、図9及び図10を用いてDBを構築する場合の処理を説明する。ここでは主に図9の上半部と図10が参照される。SAN-Mは、RAID-A内の個々のボリュームを生成 (一部システム更新) し、RAID-Aは、生成が終了するとシステムを更新 (システム管理情報、システム障害情報等) をしてRAID-DBを更新し、その更新が完了すると、SAN-MにRAID-DB構築完了の応答をする。SAN-Mは、その応答後、RAID-DBから必要なシステム管理情報とシステム障害情報を要求してから取得し、M-DBを更新する。

【 0 0 3 3 】

SAN-Mは、M-DBの更新が終了すると、SAN-FMに対してFM-DBの更新を要求し、SAN-FMは、その要求に対して、SAN-Mにシステム管理情報の送信を要求する。SAN-Mは、その要求に対して、SAN-FMにシステム管理情報を送信し、SAN-FSは、その情報でFS-DBを更新し、FS-DBの構築が終了すると、SAN-Mに構築完了の応答し、その後、SAN-FSは、その情報から接続RAID内の全ボリュームにアクセスし、SANファイルディレクトリの詳細な情報を取得し、FS-DBで管理する。ホストは、

この間のSAN-M、RAID、SAN-FSへのアクセスはロック状態（禁止）となる。

【0034】

データの転送は、信頼性を高めるために、場合により1相、2相、3相のコミット処理を行っている。ファイル記憶管理システムDB構築は、RAID単位に処理が必要となり、1度設定された状態でのリフレッシュは更新分だけとなる（1度DBが出来ればあとは差分だけの更新をする）。ファイル記憶管理システムDB構築のタイミングは、SANの管理者（オペレータ）が任意に設定が可能で、定期的なスケジューリングで更新を自動化することもできる。

【0035】

次に図10を用いてDB構築の処理の流れを詳細に説明する。ここではRAID-Aについてのみ説明するが、RAID-Bについても同様の処理が行なわれる。また、ホストはホスト-Aについてのみ言及しているがホスト-Bについても同様である。

【0036】

SAN-Mがデータベースを作る指揮権を持っている。SAN-Mはロック処理をRAID-A、SAN-FMに対して行なう。これに対して、RAID-A、SAN-FMはそれぞれロック処理を行ないロック完了をSAN-Mに返す。これ以降ホスト-Aはロックが解除されるまでRAID-Aのファイルに操作が出来ない状態になる。

【0037】

次に、SAN-MはRAID-Aに対してリフレッシュ要求を出す。リフレッシュとはデータの更新をすることである。RAID-Aでリフレッシュが終わるとリフレッシュ完了を示す情報をSAN-Mに返す。SAN-Mはこれを受けて、情報取得要求を出す。RAID-Aはこれに応答して、危機／障害／構成管理情報を送信する。SAN-Mはこの情報によりM-DBを構築する。

【0038】

SAN-Mが構成情報要求することの要求を出すとこれをトリガとしてSAN-FMが構成情報要求を出す。これに応答して、SAN-Mは構成情報を送信す

る。これによって、SAN-FMはFM-DBの構築を行なう。その後、SAN-FMとRAID-Aとの間で各LUN内のSANファイルディレクトリの取得が行なわれる。RAID-A内にはLUN0からLUNnまでであるからここでこれらを順に指定してSAN-FMに繰り返して取得する。そして、SAN-FMはその後構成情報リフレッシュ完了をSAN-Mに通知する。すると、SAN-MはRAID-A、SAN-FMにロック解除処理を行なう。これに応答してRAID-A、SAN-FMはロック解放完了をSAN-Mに返す。これによりホスト-Aはファイルの操作が出来る状態に戻る。このようにして図9のFM-DBにはRAID-Aの論理ボリュームの構成情報のDBが出来る。

【0039】

次に、ホスト-Aがファイル操作処理1をする。これについては図11、12を用いて説明する。そして、また前述のようなRAID-Aに関するDBの構築処理がなされる。

【0040】

図11及び図12を用いて、ファイル記憶管理システムDB構築の完了した状態で、RAID-Aで管理されているAファイルをオペレータAが、ホスト-Aを使って参照/更新するファイル操作の概要を説明する。

【0041】

ホスト-AからオペレータAが、任意で決定したグローバルオーナー名(C-1-15)とグローバルオーナー・グループ名(C-1-16)でSAN-FMにエントリーすると、SAN-FMは、受信したグローバルオーナー名とグローバルオーナー・グループ名をチェックし、問題がなければ、そのIDと、内部鍵と内部処理で唯一のSAN参加ID(K-1)を生成し、ホスト-Aに送信する。SAN参加ID(K-1)は、SAN-FM内部で利用されるもので、ホスト-A、オペレータAには公開しない。オペレータAは、取得したSAN参加IDをSAN-FMに送信して自分の管理しているシステム管理情報を取得し、G-DBで管理する。オペレータAは、その取得した情報からAファイルの存在とルート情報を知ることが可能で、SAN-FMにAファイル操作要求をAファイル名(K-2)をSAN参加IDに付加して依頼する。SAN-FMは、それを受信しAフ

ファイルの操作に問題がなければ、Aファイルのアクセス権、セキュリティ、識別鍵等に利用するAファイル操作ID(K-3)を生成する。

【0042】

次にSAN-FMは、RAID-AにAファイル操作ID(K-3)を付加してAファイルのステージング処理を要求する。RAID-Aは、その命令でAファイルを実空間(G-8)から仮想空間(G-7)に移動し、そのAファイルにAファイル操作ID(K-3)を付加する(ステージング処理)。ステージングされたAファイルは、SAN参加ID(K-1)で暗号化されて(暗号化されていない場合は、STR-Cで暗号化することが可能である)、Aファイルが新規の場合は仮の空間を割当てて、それを行う。また、このSAN参加IDをホストとRAID間で持ち、独自のファイルフォーマットでアクセスすることにより、ホストとRAID間通信で、改ざんやクラッカーによるデータ操作、データのすり替え等のネットチェックすることができる。

【0043】

RAID-Aは、そのステージング処理に問題がなければ、SAN-FMにAファイルアクセス許可応答をする。SAN-FMは、RAID-Aからの応答に問題がなければ、ホスト-AにAファイル操作ID(K-3)を送信し、障害があればその情報をホスト-Aに送信する。オペレータAは、ルート情報と受信したAファイル操作ID(K-3)でRAID-AにAファイル操作要求を送信し、仮想空間に割当てられたAファイルを取得する。取得したAファイルは、SAN参加ID(K-1)でネットチェックをして、型変換を行い、ホスト-Aの固有のファイルシステム(FS)で操作できるファイルフォーマットに変換/復号する。変換とは図12で多角形で囲まれた丸で示したSAN形式のファイルディレクトリが被せられていたファイルを図12で丸だけで示したホスト-Aに固有な形式のファイルに戻すことを意味する。このようにすることにより、ホスト-Aのアプリケーションプログラムがファイルを操作できる状態になる。

【0044】

Aファイルデータの操作(参照/更新)処理は、ホストとRAID間で行う。オペレータAが、Aファイルの操作を終了する時は、SAN-FMにSAN参加

I D (K - 1) と A ファイル操作 I D (K - 3) を付加して A ファイル操作終了要求を送信する。S A N - F M は、その I D をチェックし、その命令で S A N 参加 I D と A ファイル操作 I D を付加し、R A I D - A にデステージング処理を要求する。R A I D - A は、その I D をチェックし、A ファイルのネットチェックをしてデステージング処理を行い、デステージング処理が正常終了すると、更新された R A I D - D B の情報を S A N - F M に送信する。S A N - F M は、その情報で F S - D B を更新し、ホスト - A にもその情報を送信する。

【 0 0 4 5 】

ホスト - A は、その情報で正常/異常終了かを判断可能で、更新情報で G - D B を更新し、オペレータ A に結果を連絡する。A ファイルの D B 更新処理は、オペレータ A の無意識/意識的な操作で、ファイルの操作中でも、R A I D - A → S A N - F S → ホスト - A の順に D B の更新とコミット処理を行う (S A N ファイルデレクトリ情報や、システム管理情報のコミット処理)。

【 0 0 4 6 】

次に、例えばホスト - A でファイル処理をした結果ファイルの容量が大きくなってそのボリュームに収まりきれなくなるいわゆる容量障害が発生した場合について説明する。

【 0 0 4 7 】

図 1 3 及び図 1 4 で A ファイルの操作中に容量障害が発生した場合の一実施例を示す。オペレータ A が、O S - S 0 の A ファイル (A File) を操作している時、O S - S 0 の File - Space が必要な領域未満になった場合、又は、デステージング処理中に格納領域で問題が発生した場合、R A I D - A は、S A N - M と S A N - F M に対して「容量障害」メッセージを発信する。S A N - F M は、「容量障害」メッセージを受信すると、オペレータ A の A ファイル操作に障害を与えない様にする為 A ファイルの更新データ (F i l e - A ログファイル) を採取する様に R A I D - A に要求し、R A I D - A は、その命令で A ファイルのロブファイル作成を開始する (例えば、この間のトランザクション処理は続行される)。S A N - F M は、R A I D - A からのログファイル作成開始メッセージを受信すると、S A N 上の全 L U N 状態をチェックして、都合の良い空 L U N を決

定して、RAID-Aにミラーリング処理(OS-S0→OS-S1)を要求する。ミラーリング処理が終わると、RAID-Aは、SAN-FMに応答する。空きのあるLUNは図6の構成情報を参照することにより求められる。具体的には、LUN毎の全容量と存在するファイルの容量の総和を知り、全容量から総和を引くことにより空き容量が分かる。この空きのあるLUN内にボリュームのコピーを作る。これがミラーリングである。

【0048】

次にSAN-FMは、RAID-Aに一貫性処理(OS-S0→OS-S2、OS-S1→OS-S3)（重複するファイルの一方のみを消去する）を要求し、一貫性処理が終わると、RAID-AはSAN-FSに応答する。（LUN間かRAID間でファイルの断片化を防止するために、自動的か故意的にボリューム編成をすることができる。）

次にSAN-FMは、AFileの回復処理を要求し、RAID-Aは、File-Aログファイルを使って、現在の状態にAファイル(A File)を回復させ、回復後は、File-Aログファイルを自動的に削除が可能である。回復処理が終わると、RAID-AはSAN-FSに応答する。File-Aの領域は拡大され確保される。

【0049】

【発明の効果】

書換え可能な記憶媒体を用いた記憶装置群に格納されたデータを管理装置が、ファイル単位に管理が可能となり、上位装置はファイル名を指定するだけで目的のファイルを操作することが可能となり、上位装置に依存することなくファイル単位のバックアップとセキュリティ管理が可能となり、かつ個々の記憶装置の存在をユーザシステムに意識させないファイル記憶管理システムが可能となる。

【図面の簡単な説明】

【図1】本発明のファイル記憶管理システムを示すブロック図である。

【図2】本発明のRAID内部のデバイス構成を示す構成図である。

【図3】本発明のSANファイルディレクトリを説明する図である。

【図4】本発明のLUN内の記憶管理を示すDB図である。

【図 5】 本発明の L U N 内の記憶管理を示すテーブル図である。

【図 6】 本発明の R A I D 内で管理する管理 D B 図である。

【図 7】 本発明のファイル記憶管理システムの機能を示すブロック図である。

【図 8】 本発明のファイル変換フローを可視化した構成図である。

【図 9】 本発明のファイル記憶管理システム D B 構築を示す構成図である。

【図 1 0】 本発明のファイル記憶管理システム D B 構築を示す流れ図である。

【図 1 1】 本発明のファイル記憶管理システムでのファイル操作を示す流れ図である。

【図 1 2】 本発明のファイル記憶管理システムでのファイル操作を示す構成図である。

【図 1 3】 本発明のファイル記憶管理システムでの容量障害処理を示す流れ図である。

【図 1 4】 本発明のファイル記憶管理システムでの容量障害処理を示す構成図である。

【符号の説明】

A-1...ホスト-A、 A-2...ホスト-B、 A-3...SAN-FM、 A-4...SAN-M、 A-5...RAID-A、 A-6...RAIDB、 A-7...ファイバスイッチ、 A-8...LAN (A-8)、 A-9...SAN、
B-1...複数の物理デバイス、 B-2...物理デバイス群、 B-4...任意のサイズの論理ボリューム、 C-1...SANファイルディレクトリ、 C-2...ファイルエリア、 C-1-1...複数のボリュームラベル、 C-1-2...所有者名、 C-1-3...利用者識別名、 C-1-4...WWn所有者名、 C-1-5...SCSI所有者名、 C-1-6...作成日時、 C-1-7...最終変更日時間、 C-1-8...最終参照日時、 C-1-9...ボリュームアトリビュート、 C-1-10...Holder Name、 C-1-11...File Name、 C-1-12...Data格納領域情報、 C-1-13...File属性、 C-1-14...ファイルアトリビュート、 C-1-15...グローバル・オーナー、 C-1-16...グローバル・グループ・オーナー、 C-1-17...SANファイルディレクトリ、
E-2...RAID固有名、 E-3...WWn所有者名、 E-4...LUN所有者名

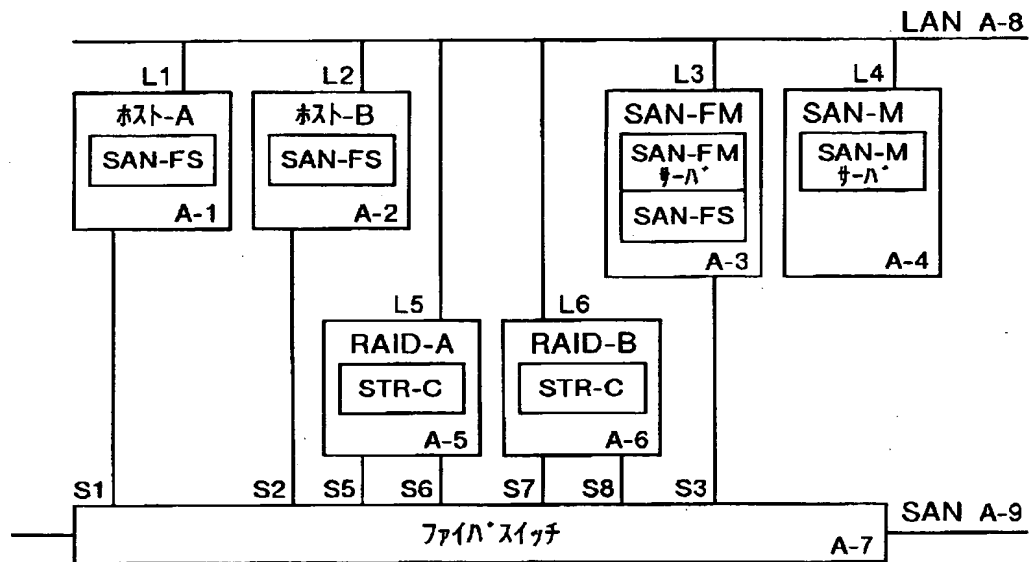
、E-5…システム管理情報、E-6…システム障害情報、
F-1、F-2…SAN-FS、F-3、F-4…STR-C、
G-1、G-2、G-3…元のファイルフォーマット、G-4、G-5、G-
6…LUNに格納されたファイル、G-7…仮想空間、G-8…実空間、
K-1…SAN参加ID、K-2…Aファイル名、K-3…Aファイル操作
ID、

【書類名】 図面

【図 1】

図 1

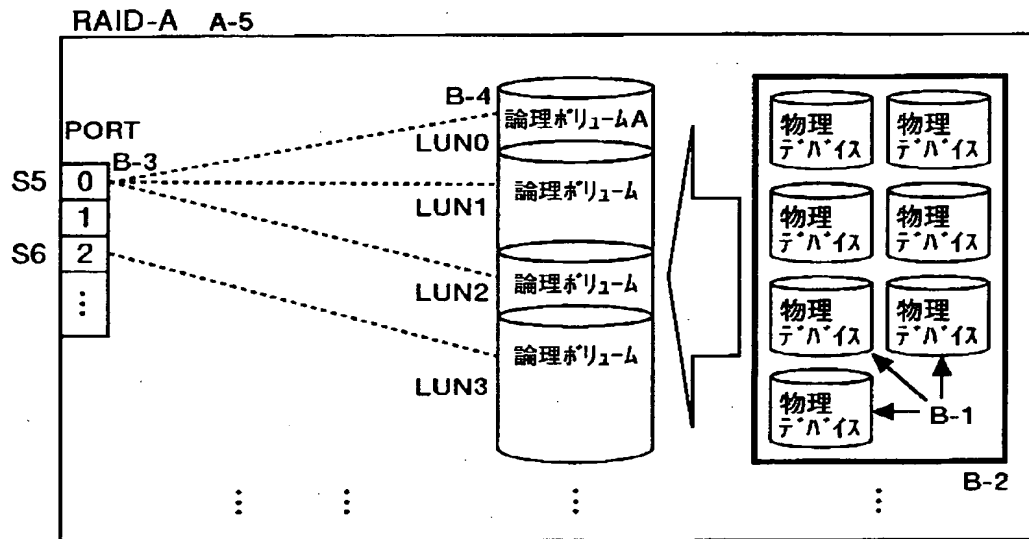
ファイル記憶管理システムを示すブロック図



【図 2】

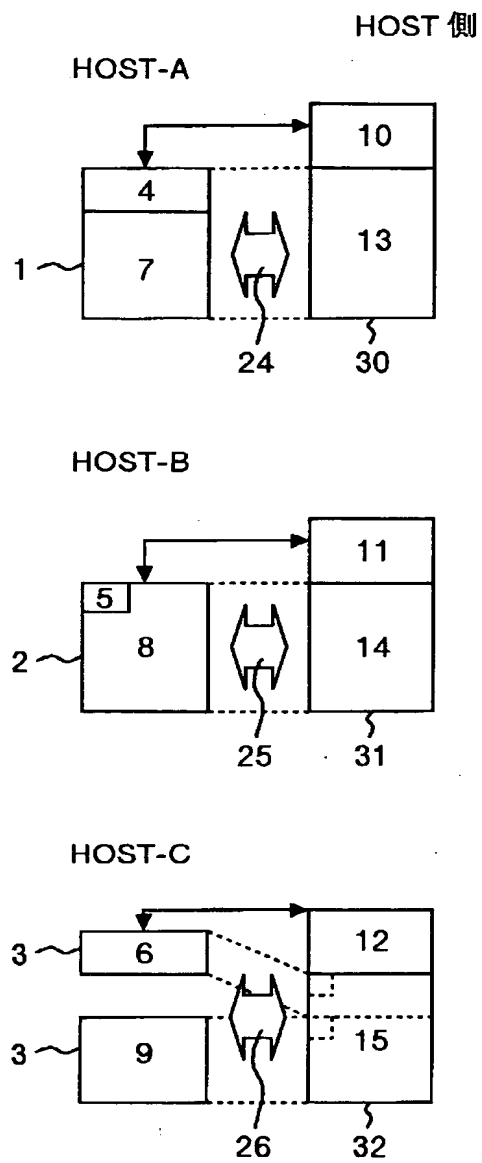
図 2

RAID 内部のデバイス構成を示す構成図



【図 3】

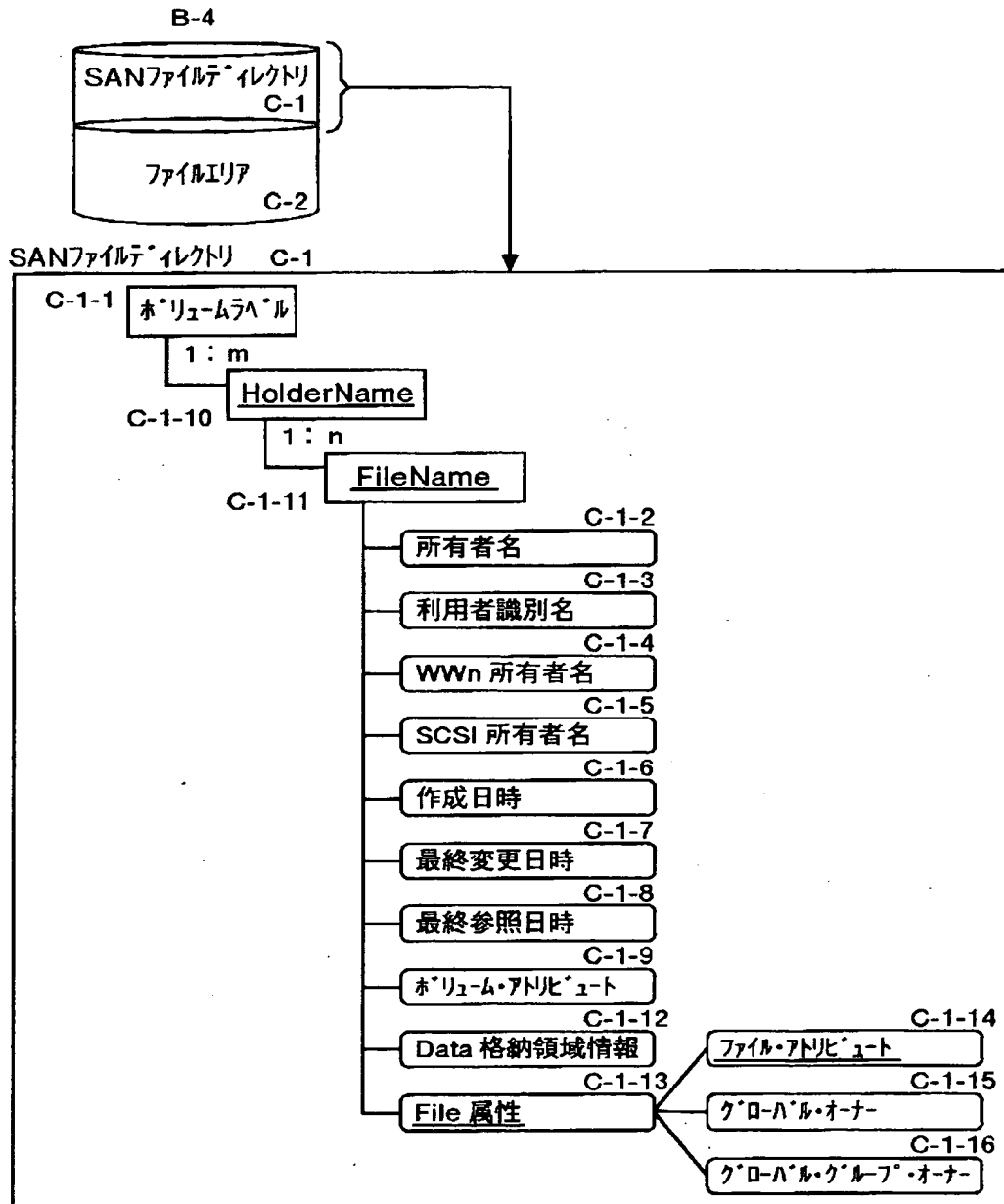
図 3



【図 4】

図 4

LUN 内の記憶管理を示す DB 図



【図 5】

図 5

LUN 内の記憶管理を示すテーブル図

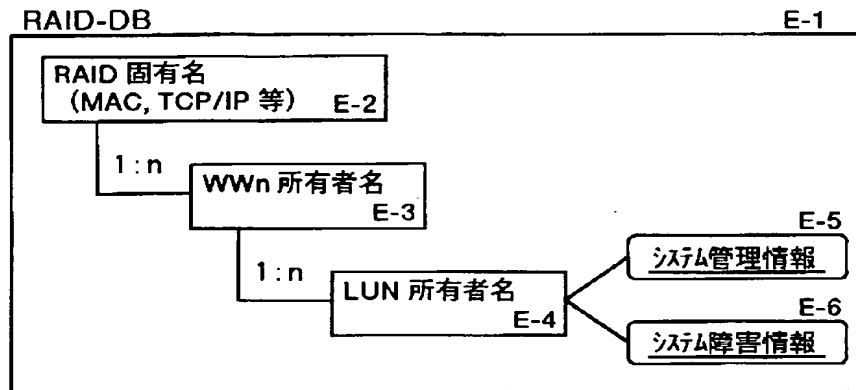
D-1 SAN7711ネットワーク (SAN-FD)

File Name	Data 格納領域情報	File 属性	ボリューム・グループ	所有者名	利用者識別名	WWn 所有者	SCSI 所有者
作成日時	最終変更日時	最終参照日時	ボリューム・グループ	HolderName			
ssssssss	hhhhhh	RRRR	VOL1	TTTT	QQQQ	0101000000	11T2L3
xxxxxxxx	xxxxxxxx	xxxxxxxx	FFFF	Holder-A			
sssssssa	hhhhha	RRRA	VOL1	TTTA	QQQA	0101000001	11T2L4
xxxxxxxx	xxxxxxxe	xxxxxxxx	FFFFFQ	Holder-B			
.	.						

【図 6】

図 6

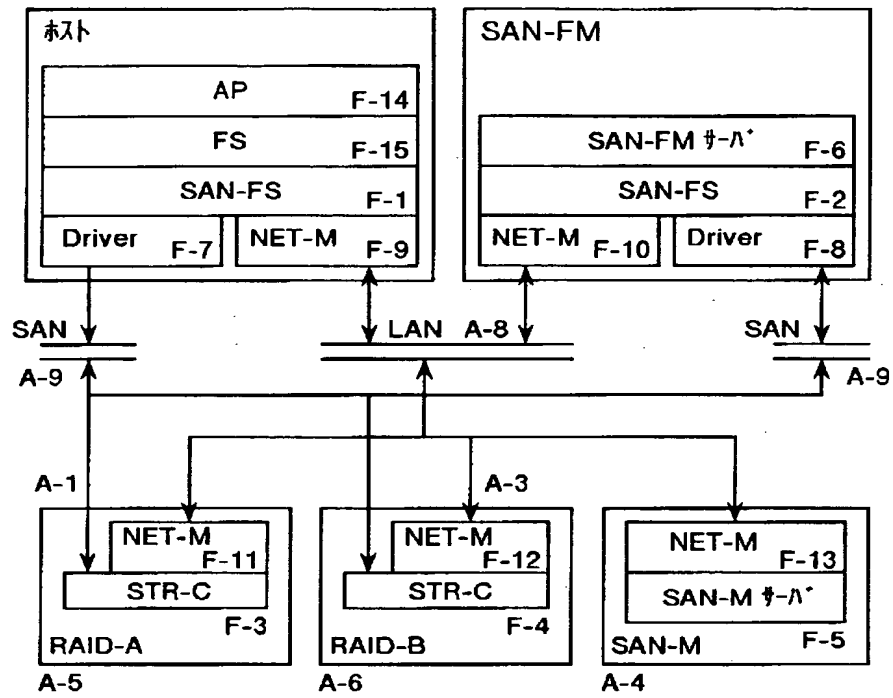
RAID 内で管理する管理 DB 図



【図 7】

図 7

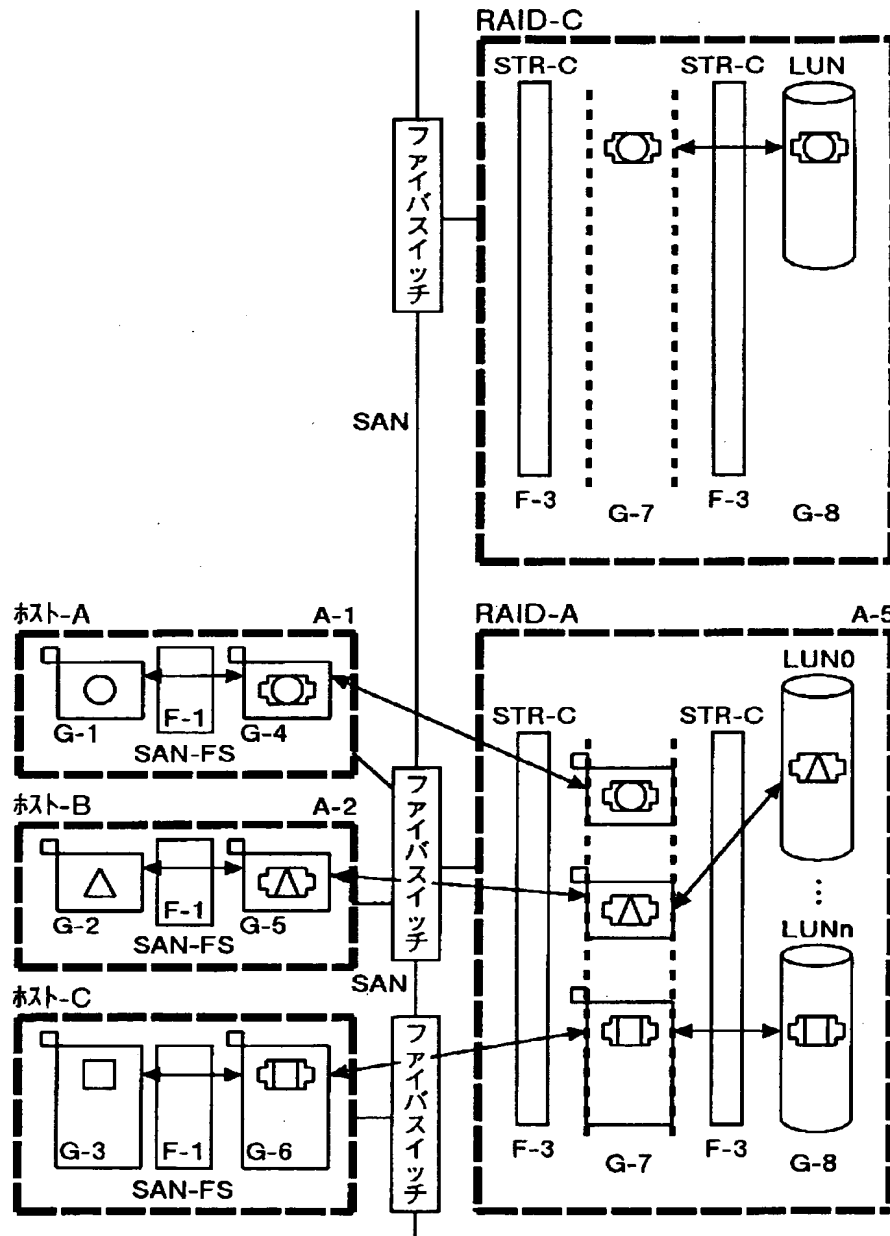
ファイル記憶管理システムの機能を示すブロック図



【図 8】

図 8

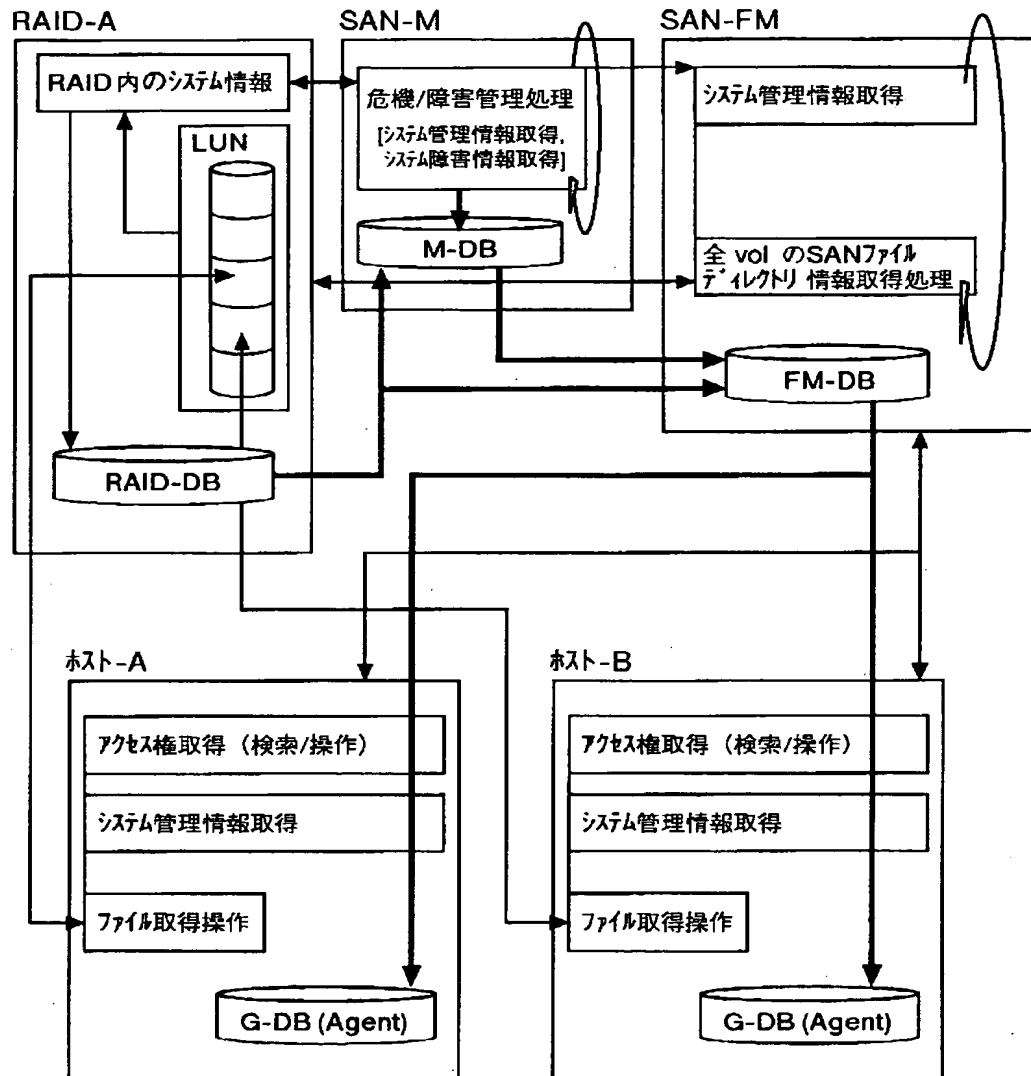
ファイル変換フローを可視化した構成図



【図 9】

図 9

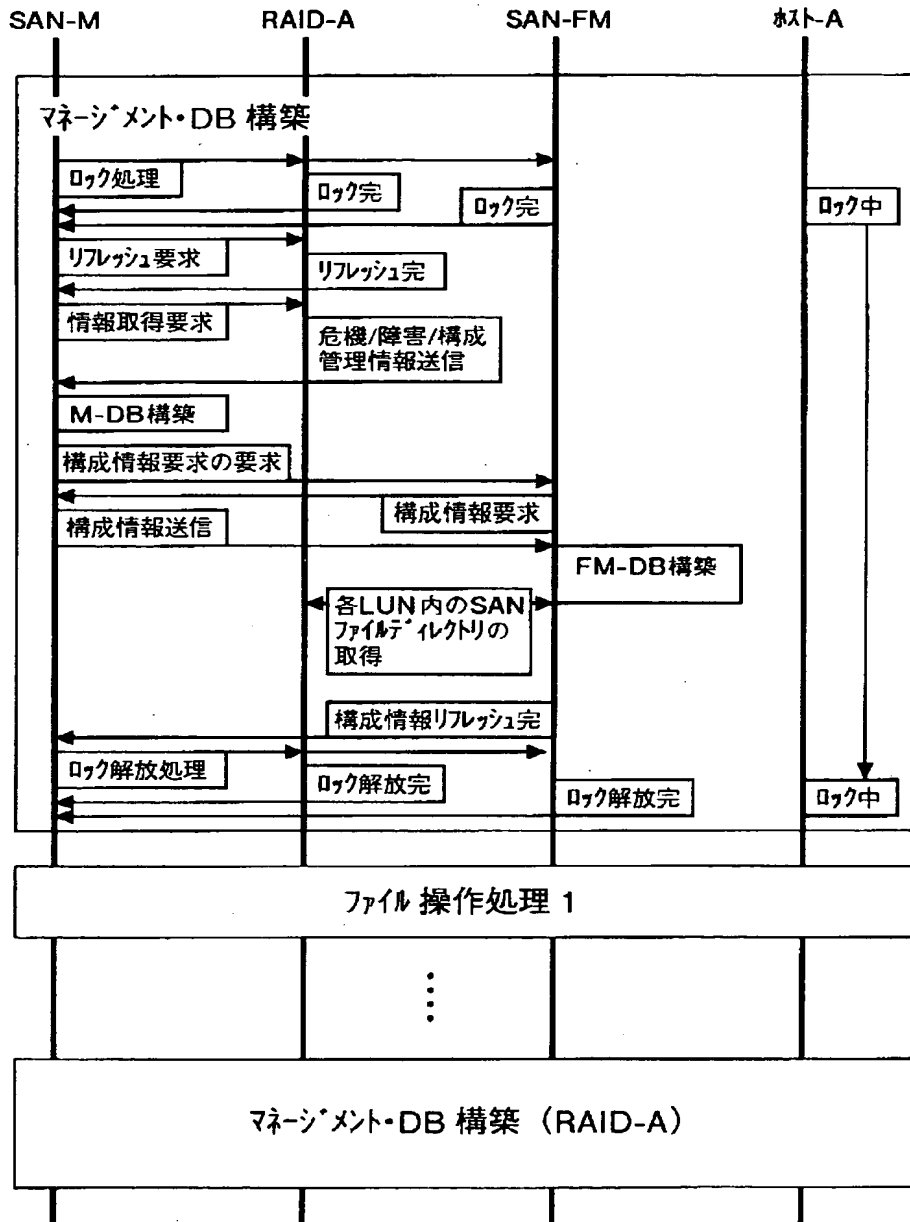
ファイル記憶管理システム DB 構築を示す構成図



【図 10】

図 10

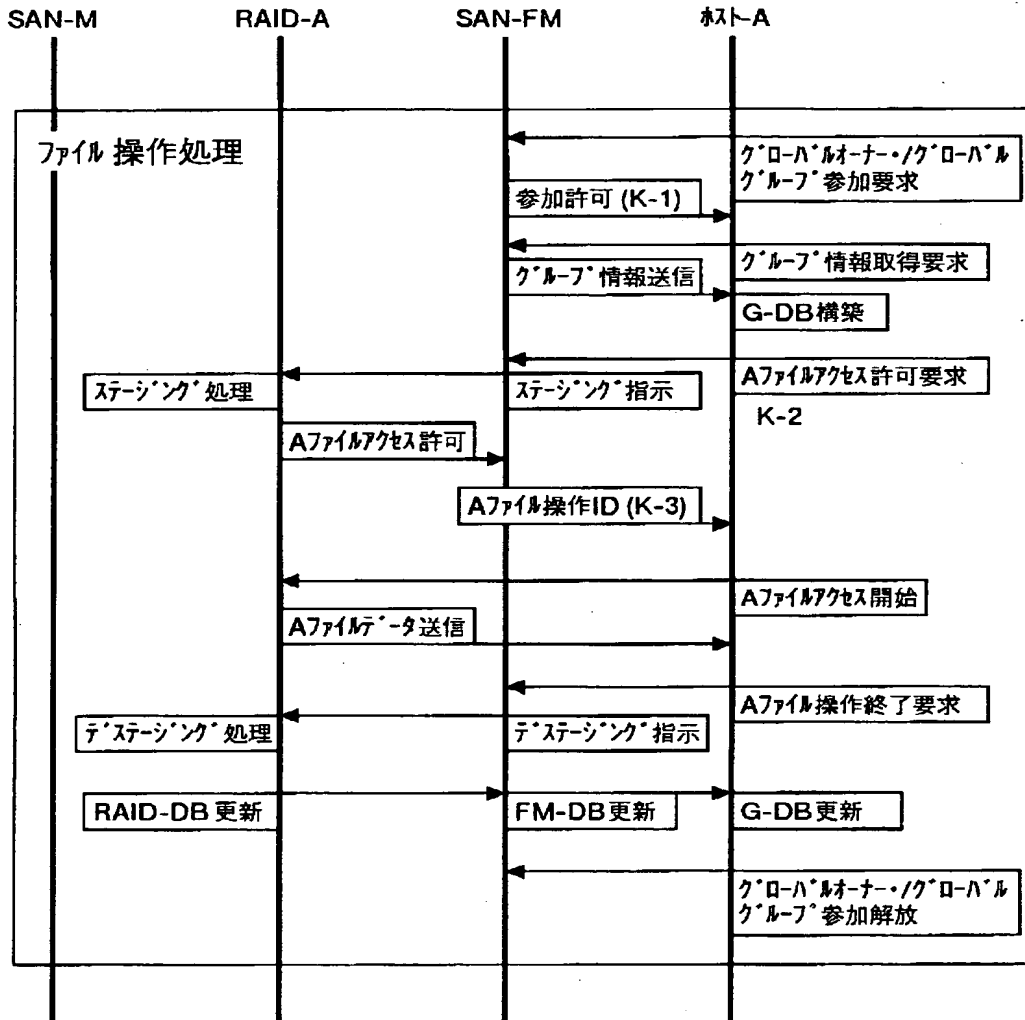
ファイル記憶管理システム DB 構築を示す流れ図



【図 1 1】

図 1 1

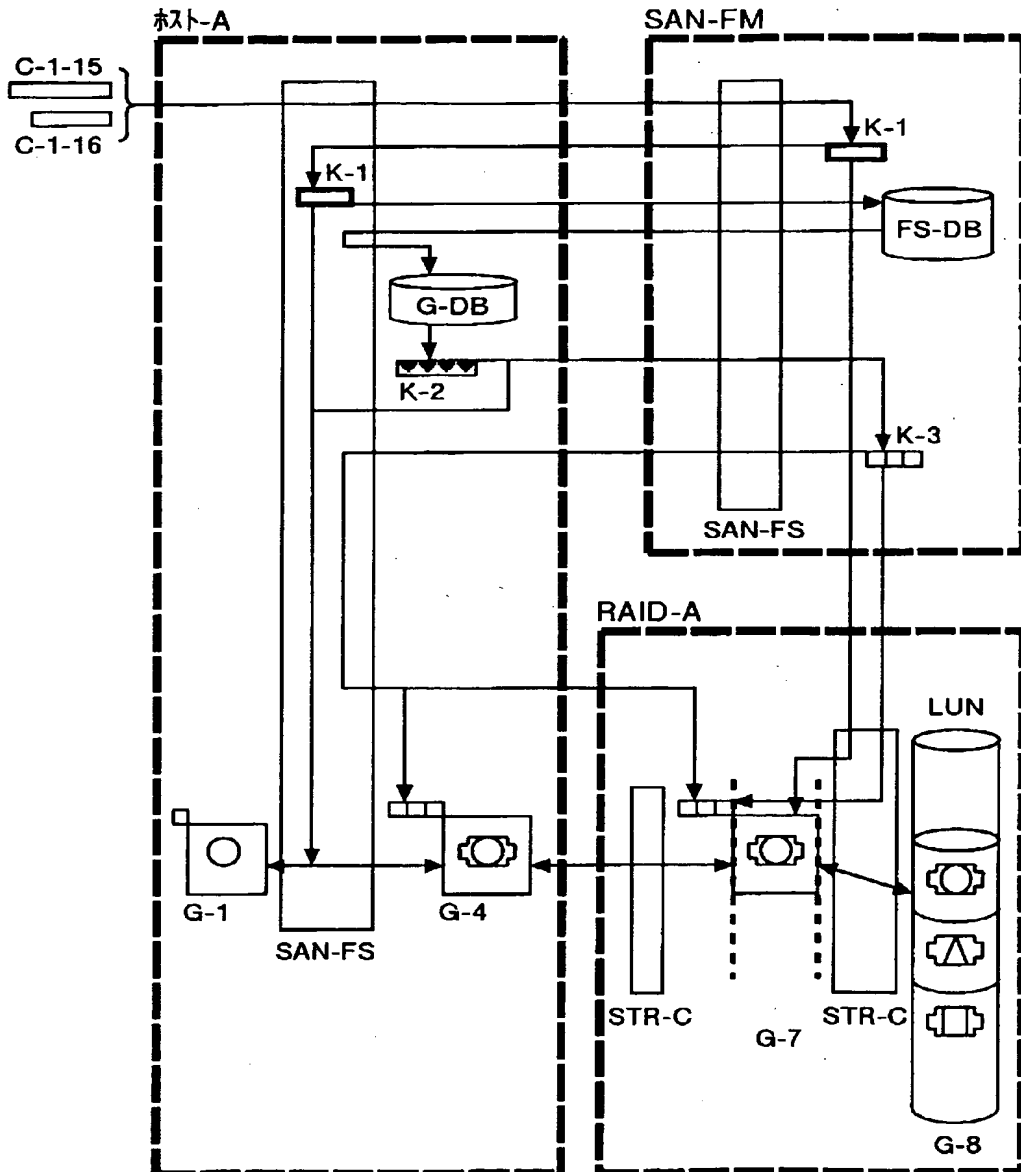
ファイル記憶管理システムでのファイル操作を示す流れ図



【図 12】

図 12

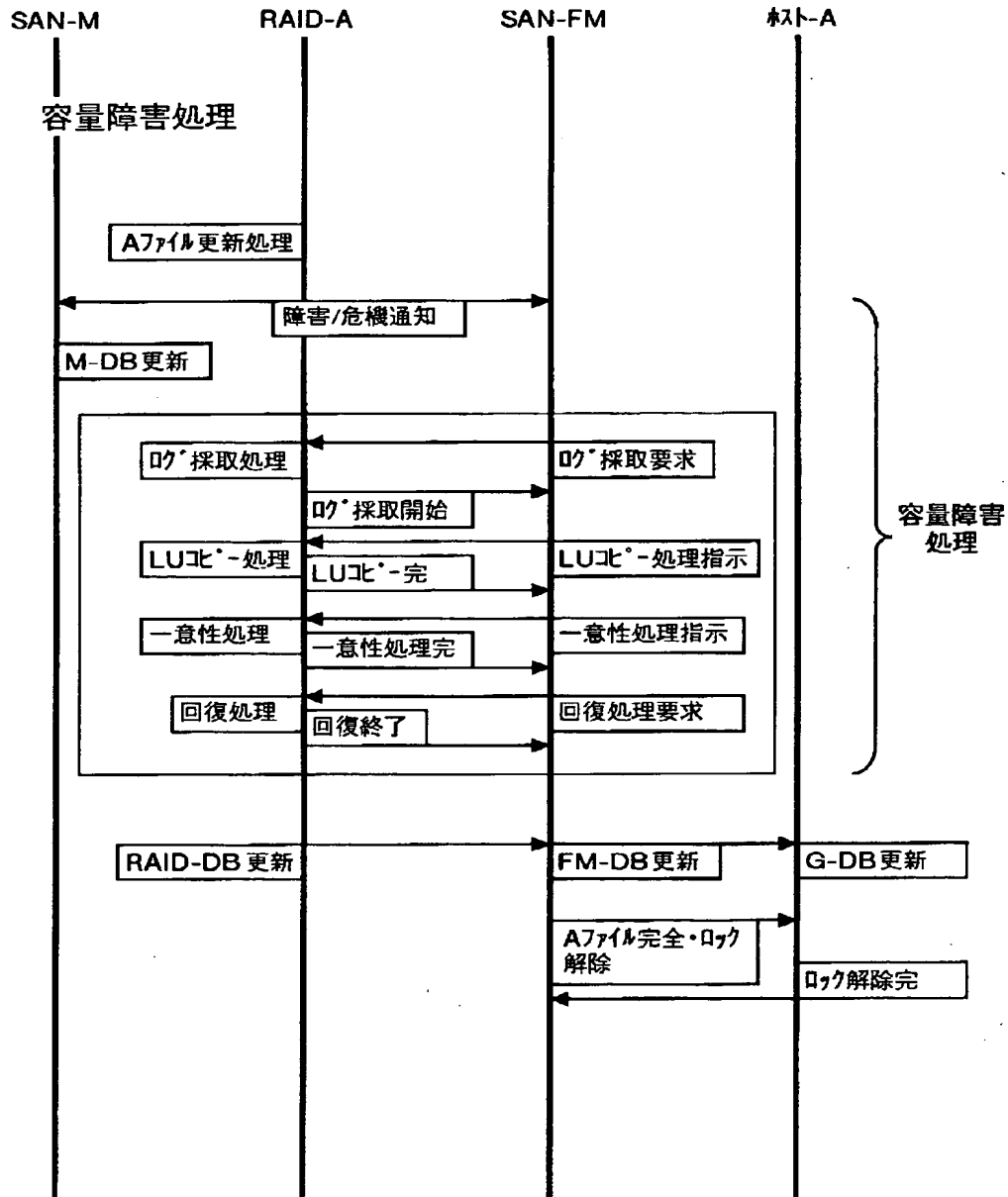
ファイル記憶管理システムでのファイル操作を示す構成図



【図 13】

図 13

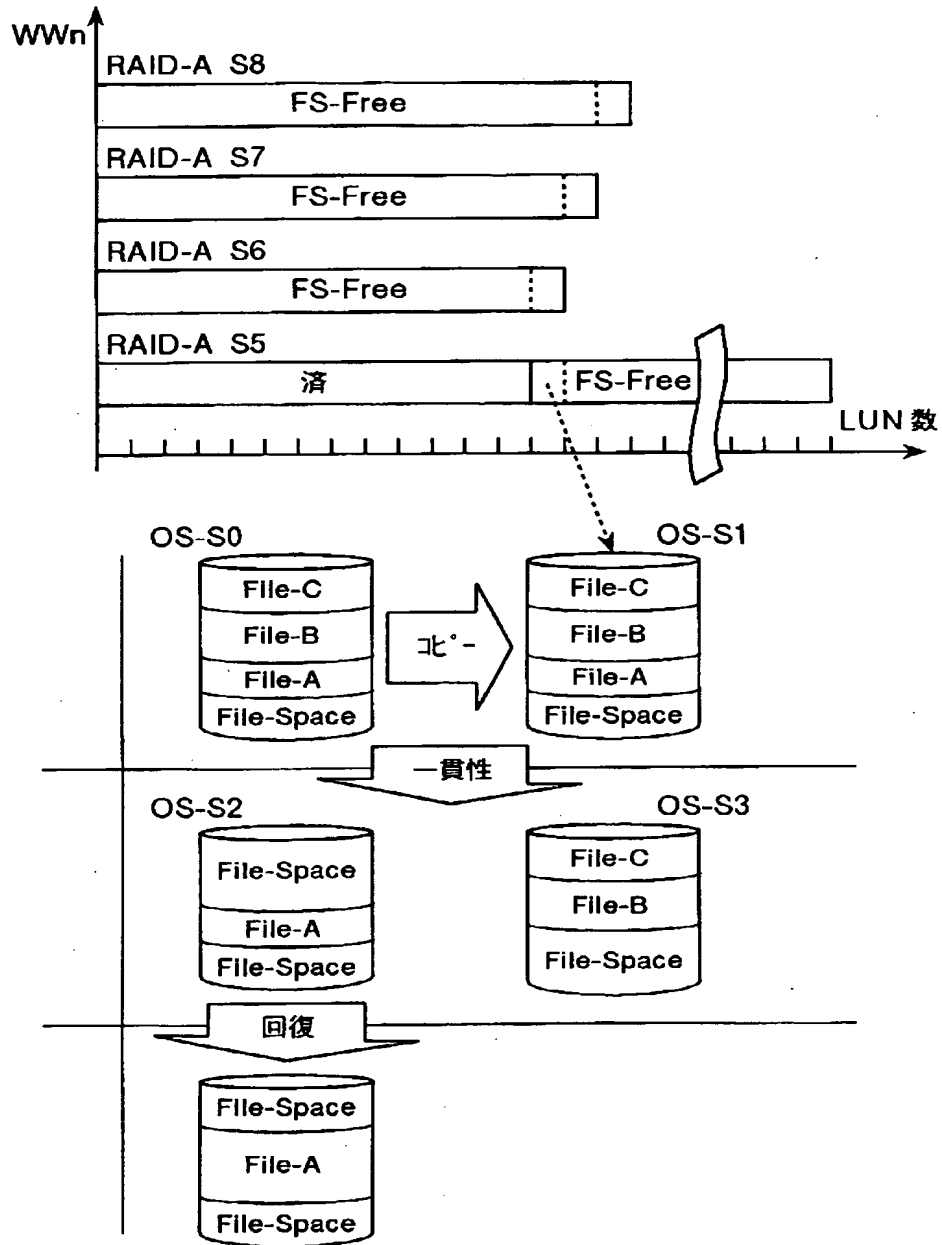
ファイル記憶管理システムでの容量障害処理を示す流れ図



【図 1 4】

図 1 4

ファイル記憶管理システムでの容量障害処理を示す構成図



【書類名】 要約書

【要約】

【課題】 S A N のような高速なデータ転送網に記憶装置が多数接続される環境にあっても、異なった O S のもとに異なったファイルフォーマットのファイルが混在していた。そのために、ストレージ管理はいつもホストの O S の管理下で実現するしかなかった。

【解決手段】 ファイルを O S に固有なフォーマットから S A N に共通なフォーマットに変換する S A N - F S を持ち、S A N - F M に記憶装置内のファイルを管理し、記憶装置間で共通のファイルでアクセス可能とする S A N - F M サーバを備えた。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日
[変更理由] 新規登録
住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.